TRANSLATOR'S DECLARATION

I, Judith Atkinson, BA. (Hons.), MITI., translator to Messrs. Taylor and Meyer of 20 Kingsmead Road, London, SW2 3JD, Great Britain, verify that I know well both the German and the English language, that I have prepared the attached English translation of 29 pages of a German Patent application in the German language with the title:

Neue für das sigM-Gen kodierende Nukleotidsequenzen

identified by the code number 000449 BT at the upper left of each page and that the attached English translation of this document is a true and correct translation of the document attached thereto to the best of my knowledge and belief.

I further declare that all statements made of my own knowledge are true and that all statements made on information and belief are believed to be true, and further that these statements are made with the knowledge that wilful false statements and the like are punishable by fine or imprisonment, or both, under 18 USC 1001, and that such false statements may jeopardize the validity of this document.

Signed: J. M. Atkinson.

Dated: 23rd January 2004

FEDERAL REPUBLIC OF GERMANY

Certificate of Priority for Filing of a Patent Application

Filing number:	100 43 337.5
Filing date:	2nd September 2000
Applicant/Proprietor:	Degussa AG, Düsseldorf/Germany First applicant: Degussa-Hüls Aktiengesellschaft, Frankfurt am Main/Germany
Title:	Novel nucleotide sequences coding for the sigM gene
IPC:	C 12 N, C 07 H, C 12 P

The attached papers are a true and accurate reproduction of the original documents for this patent application.

Munich, 21st June 2001

On behalf of the President of the German Patent and Trade Mark Office

(signature)

Faust

Novel nucleotide sequences coding for the sigM gene

The invention provides nucleotide sequences from coryneform bacteria coding for the sigM gene, and a process for the production of amino acids by fermentation using bacteria in which the sigM gene is enhanced.

Prior art

L-amino acids are used in human medicine and in the pharmaceuticals industry, in the foodstuffs industry and, very especially, in the feeding of animals.

It is known that amino acids are produced by fermentation of strains of coryneform bacteria, especially Corynebacterium glutamicum. Because of their great importance, attempts are continuously being made to improve the production processes. Improvements to the processes may concern measures relating to the fermentation, such as, for example, stirring and oxygen supply, or the composition of the nutrient media, such as, for example, the sugar concentration during the fermentation, or working up to the product form by, for example, ion-exchange chromatography, or the intrinsic performance properties of the microorganism itself.

In order to improve the performance properties of such microorganisms, methods of mutagenesis, selection and mutant selection are employed. Such methods yield strains which are resistant to antimetabolites or are auxotrophic for metabolites that are important in terms of regulation, and which produce amino acids.

For a number of years, methods of recombinant DNA technology have also been used for improving the strain of L-amino acid-producing strains of Corynebacterium, by amplifying individual amino acid biosynthesis genes and studying the effect on amino acid production.

Object of the invention

The inventors have set themselves the object of providing novel measures for the improved production of amino acids by fermentation.

Description of the invention

Where L-amino acids or amino acids are mentioned hereinbelow, they are to be understood as meaning one or more amino acids, including their salts, selected from the group L-asparagine, L-threonine, L-serine, L-glutamate, L-glycine, L-alanine, L-cysteine, L-valine, L-methionine, L-isoleucine, L-leucine, L-tyrosine, L-phenylalanine, L-histidine, L-lysine, L-tryptophan and L-arginine. Lysine is particularly preferred.

The invention provides an isolated polynucleotide from coryneform bacteria, containing a polynucleotide sequence coding for the sigM gene, selected from the group

a) polynucleotide that is at least 70% identical with a polynucleotide that codes for a polypeptide containing the amino acid sequence of SEQ ID No. 2,

b) polynucleotide that codes for a polypeptide containing an amino acid sequence that is at least 70% identical with the amino acid sequence of SEQ ID No. 2,

c) polynucleotide that is complementary to the polynucleotides of a) or b), and

d) polynucleotide containing at least 15 consecutive nucleotides of the polynucleotide sequence of a), b) or c),

the polypeptide preferably exhibiting the activity of sigma factor M.

The invention also provides the above-mentioned polynucleotide, it preferably being a replicatable DNA containing:

(i) the nucleotide sequence shown in SEQ ID No. 1, or

(ii) at least one sequence that corresponds to sequence (i) within the region of the degeneracy of the genetic code, or

(iii) at least one sequence that hybridizes with the sequence that is complementary to sequence (i) or (ii), and optionally

(iv) sense mutations in (i) that are neutral in terms of function.

The invention also provides

a replicatable polynucleotide, especially DNA, containing the nucleotide sequence as shown in SEQ ID No. 1;

a polynucleotide that codes for a polypeptide containing the amino acid sequence as shown in SEQ ID No. 2;

a vector containing the polynucleotide according to the invention, especially a shuttle vector or plasmid vector, and

coryneform bacteria which contain the vector or in which the sigM gene has been enhanced.

The invention also provides polynucleotides consisting essentially of a polynucleotide sequence, which are obtainable by screening, by means of hybridization, a corresponding gene library of a coryneform bacteria that contains the complete gene or parts thereof, using a probe containing the sequence of the polynucleotide of the invention according to SEQ ID No. 1 or a fragment thereof, and isolating the mentioned polynucleotide sequence.

Polynucleotides that contain the sequences of the invention are suitable as hybridization probes for RNA, cDNA and DNA, in order to isolate in their complete length nucleic acids or polynucleotides or genes that code for sigma factor M, or in order to isolate nucleic acids or polynucleotides or genes that are very similar to the sequence of the sigM gene.

Polynucleotides that contain the sequences of the invention are also suitable as primers, with the aid of which it is possible, by means of the polymerase chain reaction (PCR), to produce DNA of genes that code for sigma factor M.

Such oligonucleotides acting as probes or primers contain at least 30, preferably at least 20, most particularly preferably at least 15, consecutive nucleotides. Also suitable are oligonucleotides having a length of at least 40 or 50 nucleotides.

"Isolated" means removed from its natural environment.

"Polynucleotide" generally refers to polyribonucleotides and polydeoxyribonucleotides, it being possible for the RNA or DNA to be unmodified or modified.

The polynucleotides of the invention include a polynucleotide according to SEQ ID No. 1 or a fragment prepared therefrom, and also polynucleotides that are at least 70%, preferably at least 80% and especially at least from 90% to 95% identical with the polynucleotide according to SEQ ID No. 1, or with a fragment prepared therefrom.

"Polypeptides" are to be understood as being peptides or proteins that contain two or more amino acids bonded *via* peptide bonds.

The polypeptides of the invention include a polypeptide according to SEQ ID No. 2, especially those having the biological activity of sigma factor M, and also those that

are at least 70%, preferably at least 80% and especially at least from 90% to 95% identical with the polypeptide according to SEQ ID No. 2 and exhibit the mentioned activity.

The invention also provides a process for the production, by fermentation, of amino acids selected from the group L-asparagine, L-threonine, L-serine, L-glutamate, L-glycine, L-alanine, L-cysteine, L-valine, L-methionine, L-isoleucine, L-leucine, L-tyrosine, L-phenylalanine, L-histidine, L-lysine, L-tryptophan and L-arginine, using coryneform bacteria which, in particular, already produce amino acids and in which the nucleotide sequences coding for the sigM gene are enhanced, especially overexpressed.

The term "enhancement" in this connection describes an increase in the intracellular activity of one or more enzymes in a microorganism that are coded for by the corresponding DNA, by, for example, increasing the number of copies of the gene or genes, using a strong promoter or using a gene that codes for a corresponding enzyme having a high level of activity, and optionally by combining those measures.

The microorganisms provided by the present invention are able to produce L-amino acids from glucose, saccharose, lactose, fructose, maltose, molasses, starch, cellulose or from glycerol and ethanol. They may be representatives of coryneform bacteria, especially of the genus Corynebacterium. In the case of the genus Corynebacterium, special mention may be made of the species Corynebacterium glutamicum, which is known to those skilled in the art for its ability to produce L-amino acids.

Suitable strains of the genus Corynebacterium, especially of the species Corynebacterium glutamicum (C. glutamicum), are especially the known wild-type strains

Corynebacterium glutamicum ATCC13032
Corynebacterium acetoglutamicum ATCC15806
Corynebacterium acetoacidophilum ATCC13870
Corynebacterium thermoaminogenes FERM BP-1539
Corynebacterium melassecola ATCC17965
Brevibacterium flavum ATCC14067
Brevibacterium lactofermentum ATCC13869 and
Brevibacterium divaricatum ATCC14020

and L-amino acid-producing mutants or strains prepared therefrom.

The inventors have succeeded in isolating the new sigM gene of C. glutamicum coding for sigma factor M.

In order to isolate the sigM gene or other genes from C. glutamicum, a gene library of that microorganism in Escherichia coli (E. coli) is first prepared. The preparation of gene libraries is described in generally known textbooks and handbooks. There may be mentioned as an example the textbook of Winnacker: Gene und Klone, Eine Einführung in die Gentechnologie (Verlag Chemie, Weinheim, Germany, 1990), or the handbook of Sambrook *et al.*: Molecular Cloning, A Laboratory Manual (Cold Spring Harbor Laboratory Press, 1989). A very well known gene library is that of the E. coli K-12 strain W3110, which has been prepared by Kohara *et al.* (Cell 50, 495-508 (1987)) in λ-vectors. Bathe *et al.* (Molecular and General Genetics, 252:255-265, 1996) describe a gene library of C. glutamicum ATCC13032, which has been prepared with the aid of the cosmid vector SuperCos I (Wahl *et al.*, 1987, Proceedings of the National Academy of Sciences USA, 84:2160-2164) in the E. coli K-12 strain NM554 (Raleigh *et al.*, 1988, Nucleic Acids Research 16:1563-1575).

Börmann *et al.* (Molecular Microbiology 6(3), 317-326) (1992)) in turn describe a gene library of C. glutamicum

ATCC13032 using the cosmid pHC79 (Hohn and Collins, Gene 11, 291-298 (1980)).

For the preparation of a gene library of C. glutamicum in E. coli it is also possible to use plasmids such as pBR322 (Bolivar, Life Sciences, 25, 807-818 (1979)) or pUC9 (Vieira *et al.*, 1982, Gene, 19:259-268). Suitable hosts are especially those E. coli strains that are restriction- and recombination-defective. An example thereof is the strain DH5αmcr, which has been described by Grant *et al.* (Proceedings of the National Academy of Sciences USA, 87 (1990) 4645-4649). The long DNA fragments cloned with the aid of cosmids can then in turn be subcloned into customary vectors suitable for sequencing and then sequenced, as is described, for example, in Sanger *et al.* (Proceedings of the National Academy of Sciences of the United States of America, 74:5463-5467, 1977).

The resulting DNA sequences can then be studied using known algorithms or sequence-analysis programs, such as, for example, that of Staden (Nucleic Acids Research 14, 217-232 (1986)), that of Marck (Nucleic Acids Research 16, 1829-1836 (1988)) or the GCG program of Butler (Methods of Biochemical Analysis 39, 74-97 (1998)).

In that manner, the novel DNA sequence of C. glutamicum coding for the gene sigM has been obtained, which sequence, as SEQ ID No. 1, forms part of the present invention. Furthermore, the amino acid sequence of the corresponding protein has been derived from the present DNA sequence using the methods described above. The resulting amino acid sequence of the sigM gene product is shown in SEQ ID No. 2.

Coding DNA sequences that result from SEQ ID No. 1 by the degeneracy of the genetic code also form part of the invention. Likewise, DNA sequences that hybridize with SEQ ID No. 1 or parts of SEQ ID No. 1 form part of the invention. Furthermore, to those skilled in the art,

conservative amino acid substitutions, such as, for example, the substitution of glycine with alanine or of aspartic acid with glutamic acid, in proteins are known as sense mutations, which do not lead to any fundamental change in the activity of the protein, that is to say are neutral in terms of function. It is also known that changes at the N- and/or C-terminus of a protein do not substantially impair its function or may even stabilise it. The person skilled in the art will find relevant information *inter alia* in Ben-Bassat *et al.* (Journal of Bacteriology 169:751-757 (1987)), in O'Regan *et al.* (Gene 77:237-251 (1989)), in Sahin-Toth *et al.* (Protein Sciences 3:240-247 (1994)), in Hochuli *et al.* (Bio/Technology 6:1321-1325 (1988)) and in known textbooks of genetics and molecular biology. Amino acid sequences that result in a corresponding manner from SEQ ID No. 2 likewise form part of the invention.

Similarly, DNA sequences that hybridize with SEQ ID No. 1 or parts of SEQ ID No. 1 form part of the invention. Finally, DNA sequences that are produced by the polymerase chain reaction (PCR) using primers that result from SEQ ID No. 1 form part of the invention. Such oligonucleotides typically have a length of at least 15 nucleotides.

The person skilled in the art will find instructions on the identification of DNA sequences by means of hybridization *inter alia* in the handbook "The DIG System Users Guide for Filter Hybridization" from Boehringer Mannheim GmbH (Mannheim, Germany, 1993) and in Liebl *et al.* (International Journal of Systematic Bacteriology (1991) 41: 255-260). The hybridization takes place under stringent conditions, that is to say there are formed only hybrids in which the probe and the target sequence, i.e. the polynucleotides treated with the probe, are at least 70% identical. It is known that the stringency of the hybridization, including the washing steps, is influenced

or determined by varying the buffer composition, the temperature and the salt concentration. The hybridization reaction is preferably carried out with relatively low stringency as compared with the washing steps (Hybaid Hybridisation Guide, Hybaid Limited, Teddington, UK, 1996).

There may be used for the hybridization reaction, for example, a 5x SSC buffer at a temperature of approximately from 50 to 68°C. In that case, probes may also hybridize with polynucleotides that are less than 70% identical with the sequence of the probe. Such hybrids are less stable and are removed by washing under stringent conditions. That may be achieved, for example, by lowering the salt concentration to 2x SSC and optionally subsequently to 0.5x SSC (The DIG System User's Guide for Filter Hybridisation, Boehringer Mannheim, Mannheim, Germany, 1995), a temperature of approximately from 50 to 68°C being set. It is optionally possible to lower the salt concentration down to 0.1x SSC. By raising the hybridization temperature stepwise from 50 to 68°C in steps of approximately from 1 to 2°C, it is possible to isolate polynucleotide fragments that are, for example, at least 70% or at least 80% or at least from 90% to 95% identical with the sequence of the probe used. Further instructions for hybridization are commercially available in the form of so-called kits (e.g. DIG Easy Hyb from Roche Diagnostics GmbH, Mannheim, Germany, Catalog No. 1603558).

The person skilled in the art will find instructions on the amplification of DNA sequences with the aid of the polymerase chain reaction (PCR) *inter alia* in the handbook of Gait: Oligonukleotide synthesis: A Practical Approach (IRL Press, Oxford, UK, 1984) and in Newton and Graham: PCR (Spektrum Akademischer Verlag, Heidelberg, Germany, 1994).

When working on the present invention it has been found that coryneform bacteria produce amino acids in an improved manner after overexpression of the sigM gene.

In order to achieve overexpression, the number of copies of the corresponding genes can be increased, or the promoter and regulation region or the ribosome binding site, which is located upstream of the structural gene, can be mutated. Expression cassettes inserted upstream of the structural gene have a similar effect. By means of inducible promoters it is additionally possible to increase the expression in the course of the production of amino acids by fermentation. Expression is also improved by measures to prolong the life of the m-RNA. Furthermore, the enzyme activity is also enhanced by preventing degradation of the enzyme protein. The genes or gene constructs may either be present in plasmids with different numbers of copies or be integrated and amplified in the chromosome. Alternatively, overexpression of the genes in question may also be achieved by changing the composition of the medium and the manner in which culturing is carried out.

The person skilled in the art will find instructions thereon *inter alia* in Martin *et al.* (Bio/Technology 5, 137-146 (1987)), in Guerrero *et al.* (Gene 138, 35-41 (1994)), Tsuchiya and Morinaga (Bio/Technology 6, 428-430 (1988)), in Eikmanns *et al.* (Gene 102, 93-98 (1991)), in European patent specification 0 472 869, in US patent 4,601,893, in Schwarzer and Pühler (Bio/Technology 9, 84-87 (1991), in Reinscheid *et al.* (Applied and Environmental Microbiology 60, 126-132 (1994)), in LaBarre *et al.* (Journal of Bacteriology 175, 1001-1007 (1993)), in patent application WO 96/15246, in Malumbres *et al.* (Gene 134, 15-24 (1993)), in Japanese Offenlegungsschrift JP-A-10-229891, in Jensen and Hammer (Biotechnology and Bioengineering 58, 191-195 (1998)), in Makrides

(Microbiological Reviews 60:512-538 (1996)) and in known textbooks of genetics and molecular biology.

For the purposes of enhancement, the sigM gene of the invention was overexpressed, for example, with the aid of episomal plasmids. Suitable plasmids are those which are replicated in coryneform bacteria. Many known plasmid vectors, such as, for example, pZ1 (Menkel *et al.*, Applied and Environmental Microbiology (1989) 64: 549-554), pEKEx1 (Eikmanns *et al.*, Gene 102:93-98 (1991)) or pHS2-1 (Sonnen *et al.*, Gene 107:69-74 (1991)), are based on the cryptic plasmids pHM1519, pBL1 or pGA1. Other plasmid vectors, such as, for example, those which are based on pCG4 (US-A 4,489,160) or pNG2 (Serwold-Davis *et al.*, FEMS Microbiology Letters 66, 119-124 (1990)) or pAG1 (US-A 5,158,891), may likewise be used.

Also suitable are those plasmid vectors with the aid of which the process of gene amplification can be applied by integration into the chromosome, as has been described, for example, by Reinscheid *et al.* (Applied and Environmental Microbiology 60, 126-132 (1994)) for the duplication or amplification of the hom-thrB operon. In that method, the complete gene is cloned into a plasmid vector that is able to replicate in a host (typically E. coli), but not in C. glutamicum. Suitable vectors are, for example, pSUP301 (Simon *et al.*, Bio/Technology 1, 784-791 (1983)), pK18mob or pK19mob (Schäfer *et al.*, Gene 145, 69-73 (1994)), pGEM-T (Promega corporation, Madison, WI, USA), pCR2.1-TOPO (Shuman (1994). Journal of Biological Chemistry 269:32678-32684; US-A 5,487,993), pCR®Blunt (Invitrogen, Groningen, Netherlands; Bernard *et al.*, Journal of Molecular Biology, 234: 534-541 (1993)), pEM1 (Schrumpf *et al.*, 1991, Journal of Bacteriology 173:4510-4516) or pBGS8 (Spratt *et al.*, 1986, Gene 41: 337-342). The plasmid vector containing the gene to be amplified is then transferred to the desired strain of C. glutamicum by conjugation or

transformation. The method of conjugation is described, for example, in Schäfer *et al.* (Applied and Environmental Microbiology 60, 756-759 (1994)). Methods of transformation are described, for example, in Thierbach *et al.* (Applied Microbiology and Biotechnology 29, 356-362 (1988)), Dunican and Shivnan (Bio/Technology 7, 1067-1070 (1989)) and Tauch *et al.* (FEMS Microbiological Letters 123, 343-347 (1994)). After homologous recombination by means of a "cross-over" occurrence, the resulting strain contains at least two copies of the gene in question.

In addition, it may be advantageous for the production of L-amino acids to enhance, especially to overexpress, in addition to the sigM gene, one or more enzymes of the biosynthesis pathway in question, of glycolysis, of the anaplerotic pathway, of the citric acid cycle, of the pentose phosphate cycle, of amino acid export, and, optionally, regulatory proteins.

Accordingly, for the production of L-amino acids, in addition to enhancing the sigM gene, one or more genes selected from the group

- the gene dapA coding for dihydrodipicolinate synthase (EP-B 0 197 335),
- the gene gap coding for glyceraldehyde-3-phosphate dehydrogenase (Eikmanns (1992), Journal of Bacteriology 174:6076-6086),
- the gene tpi coding for triose phosphate isomerase (Eikmanns (1992), Journal of Bacteriology 174:6076-6086),
- the gene pgk coding for 3-phosphoglycerate kinase (Eikmanns (1992), Journal of Bacteriology 174:6076-6086),
- the gene zwf coding for glucose-6-phosphate dehydrogenase (JP-A-09224661),

- the gene pyc coding for pyruvate carboxylase (DE-A-198 31 609),
- the gene mqo coding for malate quinone oxidoreductase (Molenaar *et al.*, European Journal of Biochemistry 254, 395-403 (1998)),
- the gene lysC coding for a feed-back resistant aspartate kinase (Accession No. P26512),
- the gene lysE coding for lysine export (DE-A-195 48 222),
- the gene hom coding for homoserine dehydrogenase (EP-A 0131171),
- the gene ilvA coding for threonine dehydratase (Möckel *et al.*, Journal of Bacteriology (1992) 8065-8072)) or the allele ilvA(Fbr) coding for a feed-back resistant threonine dehydratase (Möckel *et al.*, (1994) Molecular Microbiology 13: 833-842),
- the gene ilvBN coding for acetohydroxy acid synthase (EP-B 0356739),
- the gene ilvD coding for dihydroxy acid hydratase (Sahm and Eggeling (1999) Applied and Environmental Microbiology 65: 1973-1979),
- the gene zwa1 coding for the Zwa1 protein (DE: 19959328.0, DSM 13115)

may be enhanced, especially overexpressed.

Furthermore, it may be advantageous for the production of L-amino acids, in addition to enhancing the sigM gene, to attenuate, especially to diminish the expression of, one or more genes selected from the group

- the gene pck coding for phosphoenol pyruvate carboxykinase (DE 199 50 409.1; DSM 13047),

- the gene pgi coding for glucose-6-phosphate isomerase (US 09/396,478; DSM 12969),
- the gene poxB coding for pyruvate oxidase (DE: 1995 1975.7; DSM 13114),
- the gene zwa2 coding for the Zwa2 protein (DE: 19959327.2, DSM 13113).

It may also be advantageous for the production of amino acids, in addition to overexpression of the sigM gene, to exclude undesired secondary reactions (Nakayama: "Breeding of Amino Acid Producing Microorganisms", in: Overproduction of Microbial Products, Krumphanzl, Sikyta, Vanek (eds.), Academic Press, London, UK, 1982).

The microorganisms produced according to the invention also form part of the invention and can be cultivated, for the purposes of the production of amino acids, continuously or discontinuously in the batch, fed batch or repeated fed batch process. A summary of known cultivation methods is described in the textbook of Chmiel (Bioprozeβtechnik 1. Einführung in die Bioverfahrenstechnik (Gustav Fischer Verlag, Stuttgart, 1991)) or in the textbook of Storhas (Bioreaktoren und periphere Einrichtungen (Vieweg Verlag, Braunschweig/Wiesbaden, 1994)).

The culture medium to be used must meet the requirements of the strains in question in a suitable manner. Descriptions of culture media for various microorganisms are to be found in the handbook "Manual of Methods for General Bacteriology" of the American Society for Bacteriology (Washington D.C., USA, 1981).

There may be used as the carbon source sugars and carbohydrates, such as, for example, glucose, saccharose, lactose, fructose, maltose, molasses, starch and cellulose, oils and fats, such as, for example, soybean oil, sunflower oil, groundnut oil and coconut oil, fatty acids, such as,

for example, palmitic acid, stearic acid and linoleic acid, alcohols, such as, for example, glycerol and ethanol, and organic acids, such as, for example, acetic acid. Those substances may be used individually or in the form of a mixture.

There may be used as the nitrogen source organic nitrogen-containing compounds, such as peptones, yeast extract, meat extract, malt extract, corn steep liquor, soybean flour and urea, or inorganic compounds, such as ammonium sulfate, ammonium chloride, ammonium phosphate, ammonium carbonate and ammonium nitrate. The nitrogen sources may be used individually or in the form of a mixture.

There may be used as the phosphorus source phosphoric acid, potassium dihydrogen phosphate or dipotassium hydrogen phosphate or the corresponding sodium-containing salts. The culture medium must also contain salts of metals, such as, for example, magnesium sulfate or iron sulfate, which are necessary for growth. Finally, essential growth substances, such as amino acids and vitamins, may be used in addition to the above-mentioned substances. Suitable precursors may also be added to the culture medium. The mentioned substances may be added to the culture in the form of a single batch, or they may be fed in in a suitable manner during the cultivation.

In order to control the pH value of the culture, basic compounds, such as sodium hydroxide, potassium hydroxide, ammonia or ammonia water, or acid compounds, such as phosphoric acid or sulfuric acid, are expediently used. In order to control the development of foam, anti-foams, such as, for example, fatty acid polyglycol esters, may be used. In order to maintain the stability of plasmids, suitable substances having a selective action, such as, for example, antibiotics, may be added to the medium. In order to maintain aerobic conditions, oxygen or gas mixtures containing oxygen, such as, for example, air, are

introduced into the culture. The temperature of the culture is normally from 20°C to 45°C and preferably from 25°C to 40°C. The culture is continued until the maximum amount of the desired product has formed. That aim is normally achieved within a period of from 10 hours to 160 hours.

Methods of determining L-amino acids are known from the prior art. The analysis may be carried out, for example, as described in Spackman *et al.* (Analytical Chemistry, 30, (1958), 1190) by ion-exchange chromatography with subsequent ninhydrin derivatization, or it may be carried out by reversed phase HPLC, as described in Lindroth *et al.* (Analytical Chemistry (1979) 51: 1167-1174).

The process of the invention is used for the production of amino acids by fermentation.

The present invention is explained in greater detail below by means of Examples.

The isolation of plasmid DNA from Escherichia coli and all techniques for restriction, Klenow and alkaline phosphatase treatment were carried out according to Sambrook *et al.* (Molecular Cloning. A Laboratory Manual (1989) Cold Spring Harbour Laboratory Press, Cold Spring Harbor, NY, USA). Methods for the transformation of Escherichia coli are also described in that handbook.

The composition of common nutrient media, such as LB or TY medium, will also be found in the handbook of Sambrook *et al.*

Example 1

Preparation of a genomic cosmid gene library from Corynebacterium glutamicum ATCC 13032

Chromosomal DNA from Corynebacterium glutamicum ATCC 13032 was isolated as described in Tauch *et al.* (1995, Plasmid

33:168-179) and partially cleaved with the restriction enzyme Sau3AI (Amersham Pharmacia, Freiburg, Germany, product description Sau3AI, Code no. 27-0913-02). The DNA fragments were dephosphorylated with shrimp alkaline phosphatase (Roche Diagnostics GmbH, Mannheim, Germany, product description SAP, Code no. 1758250). The DNA of the cosmid vector SuperCos1 (Wahl *et al.* (1987) Proceedings of the National Academy of Sciences USA 84:2160-2164), obtained from Stratagene (La Jolla, USA, product description SuperCos1 Cosmid Vektor Kit, Code no. 251301), was cleaved with the restriction enzyme XbaI (Amersham Pharmacia, Freiburg, Germany, product description XbaI, Code no. 27-0948-02) and likewise dephosphorylated with shrimp alkaline phosphatase.

The cosmid DNA was then cleaved with the restriction enzyme BamHI (Amersham Pharmacia, Freiburg, Germany, product description BamHI, Code no. 27-0868-04). The cosmid DNA so treated was mixed with the treated ATCC13032 DNA, and the batch was treated with T4-DNA ligase (Amersham Pharmacia, Freiburg, Germany, product description T4-DNA ligase, Code no. 27-0870-04). The ligation mixture was then packed in phages with the aid of Gigapack II XL Packing Extract (Stratagene, La Jolla, USA, product description Gigapack II XL Packing Extract, Code no. 200217).

For infection of E. coli strain NM554 (Raleigh *et al.* 1988, Nucleic Acid Research 16:1563-1575), the cells were taken up in 10 mM $MgSO_4$ and mixed with an aliquot of the phage suspension. Infection and titration of the cosmid library were carried out as described in Sambrook *et al.* (1989, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor), the cells being plated out on LB agar (Lennox, 1955, Virology, 1:190) with 100 mg/l ampicillin. After incubation overnight at 37°C, recombinant individual clones were selected.

Example 2

Isolation and sequencing of the sigM gene

The cosmid DNA of an individual colony was isolated using the Qiaprep Spin Miniprep Kit (Product No. 27106, Qiagen, Hilden, Germany) according to the manufacturer's instructions, and partially cleaved with the restriction enzyme Sau3AI (Amersham Pharmacia, Freiburg, Germany, product description Sau3AI, Product No. 27-0913-02). The DNA fragments were dephosphorylated with shrimp alkaline phosphatase (Roche Diagnostics GmbH, Mannheim, Germany, product description SAP, Product No. 1758250). After separation by gel electrophoresis, cosmid fragments having a size in the range from 1500 to 2000 bp were isolated using the QiaExII Gel Extraction Kit (Product No. 20021, Qiagen, Hilden, Germany).

The DNA of sequencing vector pZero-1, obtained from Invitrogen (Groningen, Netherlands, product description Zero Background Cloning Kit, Product No. K2500-01), was cleaved with the restriction enzyme BamHI (Amersham Pharmacia, Freiburg, Germany, product description BamHI, Product No. 27-0868-04). Ligation of the cosmid fragments into the sequencing vector pZero-1 was carried out as described by Sambrook *et al.* (1989, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor), the DNA mixture being incubated overnight with T4 ligase (Pharmacia Biotech, Freiburg, Germany). The ligation mixture was then electroporated into E. coli strain DH5αMCR (Grant, 1990, Proceedings of the National Academy of Sciences U.S.A., 87:4645-4649) (Tauch *et al.* 1994, FEMS Microbiol Letters, 123:343-347) and plated out on LB agar (Lennox, 1955, Virology, 1:190) with 50 mg/l Zeocin.

Plasmid preparation of the recombinant clones was carried out using the Biorobot 9600 (Product No. 900200, Qiagen, Hilden, Germany). Sequencing was effected by the dideoxy chain termination method of Sanger *et al.* (1977, Proceedings of the National Academy of Sciences U.S.A.,

74:5463-5467) with modifications according to Zimmermann *et al.* (1990, Nucleic Acids Research, 18:1067). The "RR dRhodamin Terminator Cycle Sequencing Kit" from PE Applied Biosystems (Product No. 403044, Weiterstadt, Germany) was used. Separation by gel electrophoresis and analysis of the sequencing reaction was carried out in a "Rotiphorese NF Acrylamid/Bisacrylamid" gel (29:1) (Product No. A124.1, Roth, Karlsruhe, Germany) using the "ABI Prism 377" sequencing device from PE Applied Biosystems (Weiterstadt, Germany).

The resulting crude sequence data were then processed using the Staden program package (1986, Nucleic Acids Research, 14:217-231) Version 97-0. The individual sequences of the pZero1 derivatives were assembled to a coherent contig. The computer-assisted coding region analysis was prepared using the program XNIP (Staden, 1986, Nucleic Acids Research, 14:217-231).

The resulting nucleotide sequence is shown in SEQ ID No. 1. Analysis of the nucleotide sequence gave an open reading frame of 675 base pairs, which is designated the sigM gene. The sigM gene codes for a protein of 224 amino acids.

SEQUENCE LISTING

<110> Degussa-Hüls AG

<120> Novel nucleotide sequences coding for the sigM gene

<130> jjjjjj BT

<140>
<141>

<160> 2

<170> PatentIn Ver. 2.1

<210> 1
<211> 1211
<212> DNA
<213> Corynebacterium glutamicum

<220>
<221> CDS
<222> (236)..(907)
<223> sigM gene

<400> 1
gacaccaatc cacagatgca gatcgctgaa gtacaacttg ttggttggta aattacgcgt 60

ttgtgattga cccccattaa ggtgcgcccg ccctcagttt cactaactga aggcgggcgt 120

ttttaattta tatatagctt cagctcacag gtattttcca gaaagaagag ccctcaaagt 180

atgtagcacc tcagcgacac ctcccacttg agtgggcgcc gagaagtatc tctca atg 238
Met
1

gaa aat ctg ccc ata cta agc cgc ata agg gat acg ggg tgt gtc cct 286
Glu Asn Leu Pro Ile Leu Ser Arg Ile Arg Asp Thr Gly Cys Val Pro
5 10 15

caa cct gcg ggg gat ctt atg aca gta ctg cct aaa aac cat gac cta 334
Gln Pro Ala Gly Asp Leu Met Thr Val Leu Pro Lys Asn His Asp Leu
20 25 30

agc gat acc caa ctc gtc aaa cag ttt ata tct ggc gac tcc agg gca 382
Ser Asp Thr Gln Leu Val Lys Gln Phe Ile Ser Gly Asp Ser Arg Ala
35 40 45

ttt tcc acc atc att cac cgc cac gaa cga cat atg atg cag gca gcc 430
Phe Ser Thr Ile Ile His Arg His Glu Arg His Met Met Gln Ala Ala
50 55 60 65

aga aaa tac ggg cgg aaa cca gaa gac gcc caa gac att ctc caa gaa 478
Arg Lys Tyr Gly Arg Lys Pro Glu Asp Ala Gln Asp Ile Leu Gln Glu
70 75 80

gct ctc ttt cgc gcc agc cga aac atg cac ctt tat aga gca gaa gca 526
Ala Leu Phe Arg Ala Ser Arg Asn Met His Leu Tyr Arg Ala Glu Ala
85 90 95

gct ctc ggc acg tgg ctc cac aaa ctt gtc ctg aat agc ggc ttc gat 574
Ala Leu Gly Thr Trp Leu His Lys Leu Val Leu Asn Ser Gly Phe Asp
100 105 110

```
tgg gct acc cac cgc tcc caa gta gaa ttc ccc atc ctt aac gaa cca   622
Trp Ala Thr His Arg Ser Gln Val Glu Phe Pro Ile Leu Asn Glu Pro
    115                 120                 125

aca atc gat tta gaa aaa gat cct cgc cta gcc acc gac ccc ttg ggc   670
Thr Ile Asp Leu Glu Lys Asp Pro Arg Leu Ala Thr Asp Pro Leu Gly
130                 135                 140                 145

tac ctc gat gtc gcc atg aca att cga tcc gcc atc gac caa tta cac   718
Tyr Leu Asp Val Ala Met Thr Ile Arg Ser Ala Ile Asp Gln Leu His
                150                 155                 160

ccc gat caa cgc atc gcc tta ata ctt gtc gac ctc ggc ggc tac acc   766
Pro Asp Gln Arg Ile Ala Leu Ile Leu Val Asp Leu Gly Gly Tyr Thr
            165                 170                 175

gta gaa gat gtg gcc gaa atc gaa gga atc aaa gta ggt acc gtt aaa   814
Val Glu Asp Val Ala Glu Ile Glu Gly Ile Lys Val Gly Thr Val Lys
        180                 185                 190

tca cgc cga ggg cgc gca cgc aaa gcg ttg cgc gcc ctt tta cat gca   862
Ser Arg Arg Gly Arg Ala Arg Lys Ala Leu Arg Ala Leu Leu His Ala
    195                 200                 205

gat ttc ttc ggg ccc gaa gat ggc tcc ata cag tgc gaa agc aac       907
Asp Phe Phe Gly Pro Glu Asp Gly Ser Ile Gln Cys Glu Ser Asn
210                 215                 220

tgatggaagt ttttcaaagt gtctgacgtt gaaaacggtg agttcacaac tagggtgaat 967

ggtgcacgtg atgctgcact tttacgttta ctactttgag ggaaacaatg tctgaagaac 1027

aatctgccgt agcaccaaag attcatgatg tcgccatcat cggctccggt ccagctggct 1087

ataccgcagc agtatatgca gcccgcgctg acctcaaccc catcatgttc gagggctatg 1147

aatacggtgg atctttgatg accactactg acgtggaaaa cttcccaggc tttgaaaagg 1207

gaat                                                              1211
```

<210> 2
<211> 224
<212> PRT
<213> Corynebacterium glutamicum

<400> 2

```
Met Glu Asn Leu Pro Ile Leu Ser Arg Ile Arg Asp Thr Gly Cys Val
  1               5                  10                  15

Pro Gln Pro Ala Gly Asp Leu Met Thr Val Leu Pro Lys Asn His Asp
            20                  25                  30

Leu Ser Asp Thr Gln Leu Val Lys Gln Phe Ile Ser Gly Asp Ser Arg
        35                  40                  45

Ala Phe Ser Thr Ile Ile His Arg His Glu Arg His Met Met Gln Ala
    50                  55                  60

Ala Arg Lys Tyr Gly Arg Lys Pro Glu Asp Ala Gln Asp Ile Leu Gln
65                  70                  75                  80
```

```
Glu Ala Leu Phe Arg Ala Ser Arg Asn Met His Leu Tyr Arg Ala Glu
                 85                  90                  95

Ala Ala Leu Gly Thr Trp Leu His Lys Leu Val Leu Asn Ser Gly Phe
            100                 105                 110

Asp Trp Ala Thr His Arg Ser Gln Val Glu Phe Pro Ile Leu Asn Glu
        115                 120                 125

Pro Thr Ile Asp Leu Glu Lys Asp Pro Arg Leu Ala Thr Asp Pro Leu
    130                 135                 140

Gly Tyr Leu Asp Val Ala Met Thr Ile Arg Ser Ala Ile Asp Gln Leu
145                 150                 155                 160

His Pro Asp Gln Arg Ile Ala Leu Ile Leu Val Asp Leu Gly Gly Tyr
                165                 170                 175

Thr Val Glu Asp Val Ala Glu Ile Glu Gly Ile Lys Val Gly Thr Val
            180                 185                 190

Lys Ser Arg Arg Gly Arg Ala Arg Lys Ala Leu Arg Ala Leu Leu His
        195                 200                 205

Ala Asp Phe Phe Gly Pro Glu Asp Gly Ser Ile Gln Cys Glu Ser Asn
    210                 215                 220
```

Patent claims

1. An isolated polynucleotide from coryneform bacteria, containing a polynucleotide sequence coding for the sigM gene, selected from the group

   a) polynucleotide that is at least 70% identical with a polynucleotide that codes for a polypeptide containing the amino acid sequence of SEQ ID No. 2,

   b) polynucleotide that codes for a polypeptide containing an amino acid sequence that is at least 70% identical with the amino acid sequence of SEQ ID No. 2,

   c) polynucleotide that is complementary to the polynucleotides of a) or b), and

   d) polynucleotide containing at least 15 consecutive nucleotides of the polynucleotide sequence of a), b) or c),

   the polypeptide preferably exhibiting the activity of sigma factor M.

2. The polynucleotide as claimed in claim 1, wherein the polynucleotide is a DNA, preferably recombinant DNA, that is replicatable in coryneform bacteria.

3. The polynucleotide as claimed in claim 1, wherein the polynucleotide is an RNA.

4. The polynucleotide as claimed in claim 2, containing the nucleic acid sequence as shown in SEQ ID No. 1.

5. The replicatable DNA as claimed in claim 2, containing

   (i) the nucleotide sequence shown in SEQ ID No. 1, or

(ii) at least one sequence that corresponds to sequence (i) within the region of the degeneracy of the genetic code, or

(iii) at least one sequence that hybridizes with the sequence that is complementary to sequence (i) or (ii), and, optionally,

(iv) sense mutations in (i) that are neutral in terms of function.

6. The replicatable DNA as claimed in claim 5, wherein the hybridization of sequence (iii) is carried out under a stringency corresponding to not more than 2x SSC.

7. The polynucleotide sequence as claimed in claim 2, which codes for a polypeptide containing the amino acid sequence shown in SEQ ID No. 2.

8. A coryneform bacterium in which the sigM gene is enhanced, especially overexpressed.

9. A process for the production of L-amino acids, especially lysine, by fermentation, which process comprises carrying out the following steps:

a) fermenting the coryneform bacteria producing the desired L-amino acid, in which bacteria at least the sigM gene or nucleotide sequences coding therefor are enhanced, especially overexpressed;

b) concentrating the L-amino acid in the medium or in the cells of the bacteria, and

c) isolating the L-amino acid.

10. The process as claimed in claim 9, wherein bacteria are used in which further genes of the biosynthesis pathway of the desired L-amino acid are additionally enhanced.

11. The process as claimed in claim 9, wherein bacteria are used in which at least some of the metabolic pathways that reduce formation of the desired L-amino acid are excluded.

12. The process as claimed in claim 9, wherein a strain transformed by a plasmid vector is used, and the plasmid vector carries the nucleotide sequence coding for the sigM gene.

13. The process as claimed in claim 9, wherein expression of the polynucleotide(s) coding for the sigM gene is enhanced, especially overexpressed.

14. The process as claimed in claim 9, wherein the regulatory properties of the polypeptide (enzyme protein) for which the polynucleotide sigM codes are increased.

15. The process as claimed in claim 9, wherein, for the production of L-amino acids, coryneform microorganisms are fermented in which, at the same time, one or more genes selected from the group

15.1 the gene dapA coding for dihydrodipicolinate synthase,

15.2 the gene gap coding for glyceraldehyde-3-phosphate dehydrogenase,

15.3 the gene tpi coding for triose phosphate isomerase,

15.4 the gene pgk coding for 3-phosphoglycerate kinase,

15.5 the gene zwf coding for glucose-6-phosphate dehydrogenase,

15.6 the gene pyc coding for pyruvate carboxylase,

15.7 the gene mqo coding for malate quinone oxidoreductase,

15.8 the gene lysC coding for a feed-back resistant aspartate kinase,

15.9 the gene lysE coding for lysine export,

15.10 the gene hom coding for homoserine dehydrogenase,

15.11 the gene ilvA coding for theonine dehydratase or the allele ilvA(Fbr) coding for a feed-back resistant threonine dehydratase,

15.12 the gene ilvBN coding for acetohydroxy acid synthase,

15.13 the gene ilvD coding for dihydroxy acid dehydratase,

15.14 the gene zwa1 coding for the Zwa1 protein

are enhanced or overexpressed.

16. The process as claimed in claim 9, wherein, for the production of L-amino acids, coryneform microorganisms are fermented in which, at the same time, one or more genes selected from the group

16.1 the gene pck coding for phosphoenol pyruvate carboxykinase,

16.2 the gene pgi coding for glucose-6-phosphate isomerase,

16.3 the gene poxB coding for pyruvate oxidase,

16.4 the gene zwa2 coding for the Zwa2 protein,

are attenuated.

17. A coryneform bacterium containing a vector that carries a polynucleotide as claimed in claim 1.

18. The process as claimed in one or more of the preceding claims, wherein microorganisms of the genus Corynebacterium are used.

19. A method of finding RNA, cDNA and DNA in order to isolate nucleic acids, or polynucleotides or genes, that code for sigma factor M or are very similar to the sequence of the sigM gene, which method comprises using the polynucleotide containing the polynucleotide sequences as claimed in claims 1, 2, 3 or 4 as hybridization probes.

Abstract

The invention relates to an isolated polynucleotide containing a polynucleotide sequence selected from the group

a) polynucleotide that is at least 70% identical with a polynucleotide that codes for a polypeptide containing the amino acid sequence of SEQ ID No. 2,

b) polynucleotide that codes for a polypeptide containing an amino acid sequence that is at least 70% identical with the amino acid sequence of SEQ ID No. 2,

c) polynucleotide that is complementary to the polynucleotides of a) or b), and

d) polynucleotide containing at least 15 consecutive nucleotides of the polynucleotide sequence of a), b) or c),

and to a process for the production of L-amino acids by fermentation using coryneform bacteria in which at least the sigM gene is present in enhanced form, and to the use of polynucleotides containing the sequences of the invention as hybridization probes.